



Grammar-Independent Characterizations

SL Suffix Substitution Closure:

A language L is Strictly Local iff there is a k such that for all strings $u_1, v_1, u_2, v_2 \in \Sigma^*$ and for all strings x of length k whenever $u_1xv_1, u_2xv_2 \in L$ then $u_1xv_2 \in L$.

LT Local Testability:

A language L is Locally Testable iff there exists k such that for all $u, v \in \Sigma^*$, if u and v have the same $k - 1$ prefix, k -long substrings, and $k - 1$ suffix then either $u, v \in L$ or $u, v \notin L$.

LTT Local Theshold Testabilty:

A language L is Locally Testable iff there exists k such that for all $u, v \in \Sigma^*$, if u and v have the same $k - 1$ prefix, $k - 1$ suffix, and the same number of occurrences of the same k -long substrings, counting up to some threshold t , then either $u, v \in L$ or $u, v \notin L$.

SP Subsequence Closure:

A language L is Strictly Piecewise iff whenever $w \in L$ every subsequence of w also belongs to L .

PT Piecewise Testability:

A language L is Piecewise Testable iff there exists k such that for all $u, v \in \Sigma^*$, if u and v have the same k -long subsequences then either $u, v \in L$ or $u, v \notin L$.

SF Aperiodicity:

A language L is Star-Free iff there exists n such that for all $x, y, z \in \Sigma^*$ and $m > n$ if $xy^n z \in L$ then $xy^m z \in L$.

Reg Nerode Theorem:

L is regular iff $|\{T_L(u) \mid u \in \Sigma^*\}|$ is finite, where for all $L \subseteq \Sigma^*, u \in \Sigma^*, T_L(u) = \{v \mid uv \in L\}$.

(See also the Myhill theorem.)

References

- [1] Robert McNaughton and Seymour Papert. *Counter-Free Automata*. MIT Press, 1971.
- [2] Wolfgang Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, chapter 7. Springer, 1997.
- [3] James Rogers and Geoffrey Pullum. Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information*, 20:329–342, 2011.
- [4] Jeffrey Heinz, Chetan Rawal, and Herbert G. Tanner. Tier-based strictly local constraints for phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 58–64, Portland, Oregon, USA, June 2011. Association for Computational Linguistics.
- [5] James Rogers, Jeffrey Heinz, Margaret Fero, Jeremy Hurst, Dakotah Lambert, and Sean Wibel. Cognitive and sub-regular complexity. In Glyn Morrill and Mark-Jan Nederhof, editors, *Formal Grammar*, volume 8036 of *Lecture Notes in Computer Science*, pages 90–108. Springer, 2013.
- [6] James Rogers and Dakotah Lambert. Some classes of sets of structures definable without quantifiers. In *Proceedings of the 16th Meeting on the Mathematics of Language*, pages 63–77, Toronto, Canada, July 2019. Association for Computational Linguistics.
- [7] Dakotah Lambert and James Rogers. Tier-based strictly local stringsets: Perspectives from model and automata theory. In *Proceedings of the Society for Computation in Linguistics*, volume 3, pages 330–337, New Orleans, Louisiana, 2020.
- [8] Dakotah Lambert, Jonathan Rawski, and Jeffrey Heinz. Typology emerges from simplicity in representations and learning. *Journal of Language Modelling*, 9(1):151–194, 2021.
- [9] Dakotah Lambert. *Unifying Classification Schemes for Languages and Processes With Attention to Locality and Relativizations Thereof*. PhD thesis, Stony Brook University, 2022.